Digitized by the Internet Archive
in 2017 with funding from
University of Alberta Libraries

https://archive.org/details/rayment2010

University of Alberta

Teaching Songs to Children: An Examination of the Literature on Song Acquisition

by

Meghan Margo Rayment

An essay submitted in partial fulfillment of the requirements
for the degree of Master of Music

in

Choral Conducting

Department of Music

Edmonton, Alberta

Spring 2010

DEPARTMENT OF MUSIC
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA

Abstract

This paper provides an overview of selected literature whose focus is an examination of factors affecting children's song acquisition. These factors include approaches to teaching a song by rote, use of vocal models, gesture and visual aids, developmental tonal aptitude of children, use of piano accompaniment, use of different harmonic and melodic pitch patterns, and unison versus individual singing. Through analysis of this literature, suggestions are given as to how song acquisition could be more effectively taught to young children in a choral setting.

Table of Contents

Section Page

I. Introduction…………………………………………………………1
II. A Look at the Research……………………………………………...2
III. Pedagogical Implications…………………………………………31
IV. Reference List…………………………………………………………..35

I. INTRODUCTION

The process of teaching songs to children, or song acquisition, is a topic of great importance to music education. The success of this process can greatly affect a young child's ability to participate fully in a choral setting, depending on how it is approached. Such conversations are significant because of the value of determining effective methods for helping young children ages five to ten learn how to sing new songs accurately. Recent research has provided a plethora of data from which music teachers can begin to determine the best pedagogical strategies for enriching children's acquisition of new song material.

The purpose of this paper is to begin to examine this body of research to provide the reader with an overview of some of the most important factors affecting children's song acquisition. Following a review of selected literature on this topic, I will conclude with a summary of methods that contribute to effective teaching of song acquisition.

As a conductor of a children's choir, song acquisition for the young voice has become an area of importance and interest to me. Teaching the notes, rhythms, and text of a song is an area of conducting that sometimes is passed over quickly when working with musically literate singers. The ability of adult choristers to sight-read in a rehearsal or the expectation that choristers should teach themselves the basics of a song before entering the first rehearsal, allows for the song acquisition process to occur very quickly and/or largely outside the rehearsal itself. However, when working with young children, song acquisition plays a major role in a majority of the rehearsals and therefore the approach to this portion of the rehearsal is significant and important. As familiarity with

the melody of a song must occur for the children before musical expression can be nurtured, I became interested in researching literature dealing with this topic.

Through my examination of the body of research in this area, I have discovered that there appears to be a variety of factors that may affect song acquisition. From the literature I have looked at, I have selected topics that have an influence on children's song acquisition and are of interest to me as a choral conductor. Some of the categories that will be examined in this paper include:

1) approaches to teaching a song by rote
2) use of vocal models
3) gesture and visual aids
4) developmental tonal aptitude of children
5) use of piano accompaniment
6) different melodic and harmonic pitch patterns
7) unison versus individual singing.

II. A LOOK AT THE RESEARCH

One factor that is often discussed in relation to song acquisition, is teaching by rote, or aurally, rather than visually. This is an important area of discussion as the ability of children to learn and remember a song is directly influenced by the way the song is taught. Rita Klinger, Patricia Shehan Campbell, and Thomas Goolsby (1998) compared two methods of rote teaching for accuracy. In this study, the immersion method (where the model sings the entire song repeatedly from the beginning to end and the children gradually learn the words, notes, and rhythm) and the phrase-by-phrase method (where the model sings each phrase individually and the children repeat each phrase before

putting all of the phrases together) were used to teach two songs. The results of each method of transmission were then compared to determine which teaching method yielded the most accurate results.

The participants for this study were 39 grade two students from two classes in the same, urban elementary school. All of the children were screened prior to the study to ensure the ability to match pitch. For the purpose of this study, two songs with similar characteristics were chosen. *All Around the Buttercup* and *Let Us Chase the Squirrel* are each sixteen beats in duration, have a range of a perfect fifth from the D 1 to the A 1, have an *abac* form, and antecedent-consequent or question and answer structure. Group one was taught one song by immersion and the other song by using the phrase-by-phrase method, while group two was taught each of the songs using the opposite method from group one. Both songs were taught one week and then tested one week later so as to occur at the same frequency as the children's music class. During the first meeting the songs were taught and in the next meeting, the songs were reviewed before each child was tested individually in a separate room. The tests were recorded and then analyzed by two trained elementary school music teachers.

The results of this study indicated that although the children adequately learned the two songs when taught by each method, there were significantly less errors made when they were taught using the immersion method. It was also discovered that the areas in which the greatest discrepancy in percentage of errors between the two methods occurred were in contour or the pattern of rising and falling of pitches within the melody, and pitch accuracy.

Although the results of this study appear to be valid, the test population was relatively small. The results could also have been influenced by the fact that the subjects were all from the same school and were taught by the same music teacher. Although this could be seen as a form of a control on the group being studied, this could also perhaps have influenced the test results. The results of the test may have differed due to the fact that the students were not used to more than one approach to learning a song by rote. Although it was stated that the teacher usually taught songs by rote using the phrase-by-phrase method, it could be that when the teacher introduced songs using a different method, the children were forced to listen more carefully and therefore made less mistakes using the immersion method.

One of the suggestions made by the researchers as to why the children reproduced the songs more accurately when taught using the immersion method was that the songs were in antecedent-consequent structure. This means that the first phrase of the song ended as if it was a question, while the second phrase of the song responded with a melody which could be interpreted as an answer to the first phrase. It was suggested that perhaps teaching the song phrase-by-phrase caused a lack of continuity and therefore the students could not recall as accurately what they were being taught. This comment implies that further study needs to be done in this area using four songs, two that are antecedent-consequent structure and two that are not. This would allow each study group to learn both a song in antecedent-consequent structure and one of another form using each teaching method. A conclusion could then be reached as to whether the results from this study were reached simply due to the structure of the song, or whether the immersion

method actually allows children to reproduce a song with more accuracy, regardless of structure.

In the above test, the results from one subject who could not match pitch were not included in the researcher's final results. Therefore, it may be beneficial to perform another test that includes the results of subjects who cannot match pitch. If included, more data could be collected on the rhythm, pitch, and contour aspects of this study to observe if they were different depending on whether children could match pitch or not. Notice could also be taken as to whether the accurate singing of any pitches occurred when such subjects were taught using each method. This could then provide valuable information for researchers looking for the best way to teach pitch matching to children.

Another common area of research in the area of song acquisition is the use of vocal models. Research is frequently being done to determine what type of model is the easiest for children to accurately respond. Georgia A. Green (1990) posed the question of whether the vocal model has an affect on the pitch-matching accuracy of children in grades one to six. Green also investigated which vocal model's timbre allowed for the most accurate pitch-matching from the subjects and whether gender or age affected the accuracy of responses to the models.

Green used a population of 282 elementary students form grades one to six in a university laboratory school. Audiotapes were made of an adult soprano, adult male tenor and a nine-year-old male's unchanged voice. The models were recorded singing the interval from G 1 to E 1 on the syllable *la*. Each subject was required to try to sing back the minor third interval after listening to each of the taped models. For every subject, the order of models presented was female, male, and then child. The subjects were then

tested using only one of the vocal models, allowing a lapse of seven days to occur between each test to ensure that tonal memory of the previous model was not relied on to respond to the second and third models. For the purpose of this study, each subject was allowed only one chance to respond to each taped model. When evaluated, if a pitch was determined to be within less than one hundred cents, or a half step, by an electronic tuner, it was considered to be accurate. Both notes for each response were averaged together to arrive at a classification of the entire two-note response as being either flat or sharp.

Green's results revealed that the different models did have an effect on the pitch-matching ability of the subjects from grades one to six. Analysis of the test results showed that the highest number of correct responses was to the child's voice, followed by the female model, and finally the male model. For the adult female model, the incorrect responses from both the males and females being tested were more often flat than sharp. Results also indicated that in response to the child model, both boys and girls sang more correct than incorrect responses, and that incorrect responses were more often sharp than flat. In response to the male model, boys sang more flat than correct or sharp responses, whereas girls still sang more correct than incorrect responses and their incorrect responses were more often flat than sharp. It was also determined that first, second and sixth grade students had the highest percentage of correct responses to the child model, while third, fourth, and fifth grade students had the highest percentage of correct responses to the female model. Third, fourth, and fifth grade students also had the highest percentage of sharp responses to the child model. In addition to this, first grade students had a total of less than 50 percent correct responses to the male and female models. The

correct responses for sixth grade students to the male model were also less than 50 percent.

Green used a collectively large body of students to determine the answer to the question of whether the vocal model affects the pitch-matching accuracy of grade one to six students. However, when that body is broken down into individual grade levels, the population in each was considerably less. Also, the results regarding pitch accuracy responses according to gender cannot be considered as conclusive because nowhere in the study does it state the number of boys and girls tested. In order to have a more accurate answer to the research question, the researchers should have tested an equal number of boys and girls. Research results would also have been strengthened if the researchers had selected subjects from different schools, rather than having all subjects from one school. This would eliminate the question of whether the results obtained were influenced by the students having the same female music teacher regularly, thus potentially leading to the low percentage of accuracy when the children were responding to the male model. I also found it interesting that the researchers came to the conclusion that because the male subjects sang more flat than correct or sharp responses when responding to the male model, while the girls did not, it was suggested that there may be a sociological issue with a young boy's ability to match pitch with a male voice. Although it is logical to assume that this may indicate a sociological issue, it would also be logical to assume that this may also indicate a physiological issue, as all of the boys shared this commonality while the girls did not. I feel that more research is needed before a conclusion can be reached in this area of the study. Pitch-matching problems for young male voices seems

to be a common observation, and therefore research needs to be done to determine exactly what causes this problem.

Another aspect of research on vocal models deals with the vocal quality of the model used. Cornelia Yarbrough, Judy Bowers, and Wilma Benson (1992) took a different approach than Green to this question by asking if vibrato affects children who sing and match pitch well versus those who do not. As with Green's study, this research also tested whether grade level or gender had an affect on the children's singing accuracy when responding to a vocal model with vibrato.

The population used for this test was 200 students from kindergarten to grade three of a university laboratory school. Prior to the commencement of the study, the children were classified as singers or non-singers through the use of a pitch-matching test done for another study in1995 by Yarbrough, Green, Benson, and Bowers. Included in this test were 91 singers who consistently matched pitch well and 109 singers who could not consistently match pitch. There were 45 kindergarten students, 50 grade one, 51 grade two, and 54 grade three students, with 95 females and 105 males included in this study.

For this research, three models were used: a child with no vibrato or 100% pitch accuracy, an adult female singing without vibrato with 100% pitch accuracy, and an adult female singing with vibrato that had 79.53% pitch accuracy. This means that 79.53% of the time the singer was producing the desired pitch, while the rest of the time other pitches were produced due to vibrato. This percentage was determined using a computer analysis program. Each child tested responded to a tape of each model singing a descending minor third from G 1 to E 1 on the syllable *la*, as was done in the study by Green mentioned earlier. The researchers presented the three models in a different order

for each subject tested to ensure that the order in which the models were presented did not affect the accuracy of the test results.

The results of this test revealed a significant difference between the number of correct responses given by those who could consistently match pitch and those who could not. There was also a significant difference in the percentage of correct responses to the child, adult female without vibrato, and adult female with vibrato models. For both accurate and inaccurate singers, there was not a significant difference between grade levels on the accuracy of their responses. When responding to the child model, females sang significantly more correct responses than males in every grade and there was no difference between grade levels.

In response to the female vibrato model, there were more correct than incorrect responses for females in the first and second grades and less correct responses for females in grade three. Contrary to this, accuracy in responses from males and females in kindergarten were fairly equal. In response to the non-vibrato adult female, there were no significant differences between males and females or between grades. As far as the order in which the models were presented, it seemed to affect only the singers who could not consistently match pitch. More accurate responses were sung to the vibrato model the later it was presented during the test. In order to obtain a 49% accuracy rate of responses to the female model with vibrato, this model had to be preceded by either the female model without vibrato or the child model. Responses to the non-vibrato model were fairly consistent at 49 to 53% correct responses, regardless of where they were placed in the test order. The child model also had to be preceded by one of the other models before attaining 49% correct responses

This test seems to have been fairly well controlled: there was a comparable number of children per grade, a comparable overall number of female and male participants, and a similar number of singers who could and could not consistently match pitch. Because of the considerations taken to ensure that results were accurate, I would tend to agree with the conclusions drawn from this one test. It was concluded that the non-vibrato female adult model seemed to elicit most consistently the most accurate responses. With the other two models, it seemed to depend on where they occurred in the performance rotation of the model as to how many accurate response were given.

Regarding the results comparing the model's effect on inaccurate and accurate singers, the accuracy depended largely on the test used to classify the subjects as being either primarily accurate or inaccurate singers. Further research would need to be done to ensure that the prior classification test was accurate, thus allowing the conclusions reached in this test to also be accurate.

For further research, it would be interesting to perform this same test using a phrase of music or short song, rather than the interval of a minor third, to test pitch accuracy. This would address more directly the question of whether a model with or without vibrato could affect contour and pitch accuracy in song acquisition, not just pitch-matching with individual intervals.

Mei-Ying Liao's study (2008) was a well-designed experiment, allowing for what seem to be quite convincing results. Although there were several questions being asked in this study, the process and conclusions were well organized and clearly presented. The questions being asked dealt with the following: whether gesture would help children sing more accurately; whether these gestures would have a more influential effect on different

melodic contours; and whether gender would affect the results of the use of gestures on vocal accuracy.

Liao used a population of 80 children between the ages of five and six. Because gender comparisons were also being made, an equal number of boys and girls were included in the study. The children used for this study were randomly selected from four different kindergartens in Chungli and Hsinchu, North Taiwan.

For the purpose of this study, the researcher decided to write new tonal patterns that fit the test requirements rather than use songs that had already been written. This was done in order to avoid inaccuracies if children were familiar with songs that were chosen. The neutral syllable *bum* was chosen for the children to sing the patterns. Liao also created the gestures the children performed while singing the test for this study. For ascending lines, the arms began in a closed position in front of the body, with the hands placed together, and were gradually opened by large or small increments depending on whether the melodic movement was in steps, small leaps, or large leaps. For descending melodic movement, the opposite was done. This opening and closing of the arms was to symbolize the opening of the throat and jaw, as well as an increase in energy needed when singing higher pitches and a smaller space needed for lower pitches. For repeated pitches, both arms were moved forward in a fluid motion. These gestures were tested for difficulty on three boys and three girls who were not included in the test population.

The children in the test population were first taught the gestures together using imitation and verbal communication. Then, prior to the test, the children followed a video recording to test whether they understood the process that would be used in the actual test. At the time of the test, the children each drew a card to determine whether they

would be tested with or without gestures first. The probability rate for the results of the draw was fifty percent. Each child was then tested individually with all six of the melodies and afterwards was asked whether he/she preferred singing with or without gestures, and why. It was not stated how many times each video recording was played or how many chances each child had to respond to the recording.

Data revealed that the children sang more accurately when using gestures than when singing without them. It was also found that the girls sang more accurately than the boys regardless of whether they were using gestures. In response to the question of whether the subjects preferred singing with or without gestures, most responded that they enjoyed using gestures.

One aspect of this study that could have been presented more clearly in this article was how the children were prompted to sing the melodic patterns individually. The article did not state whether the video was used or whether a live model was used to sing the patterns prior to the children singing each pattern by themselves.

I feel that this study was well organized and that the results seem to be quite accurate. Liao used an equal number of boys and girls because gender issues were being compared. This is an important feature in order to get accurate results, and it is one that is not always considered by researchers when comparing gender issues. Liao also classified the children by age and by grade level. This is important for accuracy because if only grade level is considered, there can be children of varying ages in each grade without the researchers knowing. This could therefore affect the accuracy of the results as a significantly older or younger child within a grade could have much different musical abilities than the rest of the children in that grade.

Due to the perceived accuracy of results found in this study, I would agree with the conclusions drawn by the researcher. Explaining the vocal mechanism and how it functions when used properly and improperly is a complex process. When working with young children, this complexity can cause confusion and allow for the possibility of incorrect or unhealthy vocal production. By using gestures, some of this confusion can be avoided and the children's focus can be on performing the gesture, rather than positioning of the vocal mechanism. When teaching children to sing, movements are something that I commonly use to allow for a healthy vocal production, without having to explain what needs to happen physiologically to achieve this. I also feel that gestures should not necessarily be fixed for a specific melodic contour. Although this was necessary for the purposes of this study, modifications may be necessary depending on the child and his/her current vocal abilities.

Hilary Apfelstadt (1984) dealt with this same topic by testing the idea of aiding children in the understanding of melodic contour by using visual and kinesthetic aids. Apfelstadt's main goal was to understand if training children in melodic perception would affect the singing accuracy of students in kindergarten. The other questions this study hoped to answer were whether music in the home affects vocal accuracy, whether a relationship exists between vocal accuracy and pitch discrimination, and whether gender affects achievement on tests of vocal accuracy and pitch discrimination.

Apfelstadt used 61 students from three kindergarten classes, with each class representing one group in the study. The classes used for this study were from public schools in Madison, Wisconsin. The first group used both visual aids and movement to help teach melodic perception. The second group was taught to sing using simple

imitation. The third group was the control group and was taught in a traditional manner using activities and encouraging participation, without the goal of teaching melodic perception. In order to test the other research questions, the students were classified by gender and questionnaires about the presence or absence of music in the home were completed by the parents. In group one there were 8 males and 16 females. Group two had 9 males and 13 females and group three had 6 males and 9 females.

Prior to the commencement of the study, Apfelstadt went into the music classes of groups one and two and taught an unspecified number of songs by rote. Following this, there was an individual pre-test to determine each child's understanding of melodic patterns through a rote singing test and a test for vocal accuracy. The latter involved having each child sing back 20 melodic patterns, each of which was played three times from a tape of a vocalist, and then sing a song of the student's choice. After these tests, the students were classified as having high, medium, or low vocal accuracy. Between the pre-test and the test, groups one and two received two 30-minute classes of instruction each week for eleven weeks. During this same time, group three received two 20-minute sessions and one thirty-minute recreational singing period each week.

After the eleven weeks of instruction, the tests for vocal accuracy and rote-singing were both given again. This time, the rote-singing test included two songs taught to all of the children.

At the conclusion of the tests, Apfelstadt suggested the data indicated that gender did not affect the results of vocal accuracy. It was also concluded from the pre-test results and parent surveys, that less accurate singers generally came from homes classified as having low to medium musical environments. More accurate singers generally came from

homes classified as having medium to high musical environments. The researcher concluded that although teaching with emphasis on melodic perception did seem to affect pitch pattern accuracy, it did not seem to significantly increase vocal accuracy or pitch discrimination. Both groups one and two were better than group three in pitch pattern singing. Also, group two improved more in accuracy of rote singing than did group one over the eleven-week training period.

I would agree with Apfelstadt that the results may not be completely accurate. One possible reason for this is that group three did not have the same teacher as groups one and two. This could have been a contributing factor to the lower test scores of group three. The average age of the students in group two was also the youngest which, as suggested by the researcher, may have led to their ability to increase their rote singing accuracy over the eleven-week training period. Results could also have been more accurate had there been an equal number of males and females included in this test. Another factor possibly affecting accuracy, was the inequality in class times in the training period. Without all three groups having the same amount of instructional class time, it is very hard to make a reliable comparison of results. I also feel that the children's sing-back test results may have been more accurate had they been given the pitch in a different way than using the pitch pipe. It would seem to make the most sense to give the first pitch to the children in a way that has been previously suggested, thereby allowing them the best opportunity to correctly match pitch. Perhaps using a child vocal model or a female vocal model would have allowed for more accurate pitch matching of the first note (Green 1990, 227; Yarbrough, Bowers, Benson 1992, 37).

Another factor affecting song acquisition is developmental tonal aptitude or an individual's ability to do well in music. Christina M Hornbach and Cynthia C. Taggart (2005) focus on three different questions involving this. The first question is whether kindergarten, grade one, two, and three students' developmental tonal music aptitude affected their achievement when singing. Also asked was whether age of the child or the school they attended affected the student's achievements.

The population tested for this study was 162 kindergarten, grade one, two, and three students from Michigan public elementary schools. The students were randomly selected, although students at both schools did have music class twice each week. It was known prior to the test that the teacher in the second school used solo singing in her classes, thus allowing the children the chance to practice singing alone as well as in a group more frequently than the teacher in the first school.

In each of the two school's regular music classes, the students were given the Tonal Subtest of the Primary Measures of Music Audiation by their music teachers. This is a test in which pictures are used to allow a child to express whether two phrases of music are the same or different (Gordon 1979, 42). Within three weeks of the first test, a second test was given to examine the children's ability to sing a piece of music in the correct key with the correct melodic contour. This test was designed by the researchers for the purpose of this study. The song used for this test was *Bow Belinda.* The test was not done until each class had worked on this song for four music classes, each of which was approximately thirty minutes in length. Each class was then divided into groups of three to five students each group of which was tested together. In each small group the students reviewed the song twice with the researchers and then were asked to sing the

song individually. Before each child sang, one of the researchers would play a G on a soprano recorder to give the first pitch of the song and then sing *sol-la-sol-fa-mi-re-ti-doh* on a neutral syllable to establish the tonal centre.

Each child's singing was marked out of five, with one being the most accurate singing and five being a performance in which the song was sung or chanted to a different melodic contour than the one taught to and reviewed with the children. The three judge's marks were then combined to arrive at a mark ranging from three to fifteen.

Data indicated that the average mark tended to rise with each grade level through to grade two, and that the achievement rate of the third-grade students was lower than the students in grade two in both schools.

Students from school two scored higher on the singing achievement test than students from school one. However, the results were quite low, regardless of which grade or school the children were in. It was also found that students in grade two had higher averages than students in kindergarten and grade one. Grade three students also had higher averages than kindergarten students on the singing achievement test, but lower than students in grade two.

It was concluded by the researchers that when singing is taught to students, their singing achievement might improve. It was also determined from the results of the test that tonal aptitude, as determined by the initial Tonal Subtest of the Primary Measures of Music Audiation test, and singing achievement are not dependent on one another.

One thing that was not clear in this article was whether the 162 randomly selected students from kindergarten to grade three were evenly divided between the four grades. If there were not an equal number of students from each grade tested, the results would not

have been as accurate as they could have been. When singing accuracy is being tested and compared by grade level, it is important to have the same number of children per test group or grade. Also, grade level only, and not age, was discussed. This again could potentially lead to inaccurate results as a child in grade two could potentially be the same age as a child in grade three, but their test results are still being considered with the remainder of the grade two students' results.

I also question the researcher's decision to give the first note for each student's test with a soprano recorder instead of singing it, as well as the tonal pattern, before each child sang. It was stated earlier that children sang more accurately when a child voice was used as a model. Nowhere in this article was it mentioned that a soprano recorder elicited a more accurate singing response than a child voice, so it was interesting that this instrument was chosen to give the first pitch

I also found it interesting that the researchers would choose to sing the pattern *sol-la-sol-fa-mi-re-ti-doh* to help the children establish a tonal centre. Although this pattern would definitely establish a key centre, the children may have been more familiar with a major scale in tonic *sol-fa*. This would allow the children to hear all of the pitches in sequence within an octave. If the researchers desired the last pitch sung to also be the starting pitch of the piece, they could simply have sung the scale ascending and descending, thus allowing the children to hear *doh* sung three times.

I did find it interesting, and not surprising, that school two had a higher average for the singing achievement test than did school one. It was stated earlier in the article that the music teacher in school two put more emphasis on solo singing in class; therefore, since the test was given in a solo singing setting, it makes sense that the

children who had more experience singing solo would achieve a higher score. It has also been suggested that children sing more accurately individually rather than in a group (1985 Goetze). The children in school two would therefore have had more experience singing individually and would therefore be able to develop the skills needed for accurate singing when learning the test song in a group. I did find it surprising that the grade three students of both schools sang with less accuracy than the grade two students. It would be very interesting to do more research on this topic to see if this result was in fact a social construct.

Another area of study affecting song acquisition is how pitches are presented -- melodically or harmonically. Melodic presentation could refer to the modes which the melodic patterns are based on, or the contour of the melody. Harmonic presentation could mean which position of the chord the target note is presented or whether there is piano accompaniment with a melodic line. Randall S. Moore asks the following three questions in his study: 1) How accurately do children sing back tones from melodies and chords they hear; 2) Which tones from those chords are the easiest for children to reproduce; and 3) If mistakes are made in trying to reproduce tones, what are they?

Moore (1994) selected for his study 128 children from 56 different schools, who had been designated by their teachers as good singers and who ranged in age from eight to eleven. The children were divided into four equal groups of 32 students each, grouped according to age and sex. Group one was made up of girls from ages eight to nine, group two was boys from ages eight to nine, group three was girls ages ten to eleven, and group four was boys ages ten to eleven.

In order to answer the research questions, each child underwent a ten-minute pitch-matching test, which consisted of four melodic and twelve harmonic patterns. The harmonic patterns were divided into four upper tone pitch matching tests made up of a diad or two-note chord, four lower tone pitch-matching tests made up of a diad, and four middle tone pitch-matching tests made up of a triad. To ensure that order did not affect accuracy of the results, half of the children did the melodic pitch-matching test first, while the other half did the harmonic pitch-matching test first. All tests were given using the piano to play each melody or chord at a tempo of quarter note equals 100 beats per minute. Each of the melodic examples used for the test progressed from short to long, and simple to difficult. All melodies and chords tested were within the range of middle C to the C one octave above. The children were asked to sing back the melodies or specified tone of a chord on the syllable *la*.

No significant difference was found when upper and lower pitch-matching tests were reversed. No significant difference was found between sexes or age groups in their pitch-matching abilities. However, it did take eight and nine year olds more repetitions to match the lower tone in a diad than it did for ten and eleven year olds. Both age groups had difficulty matching the middle tone of a triad. It was discovered that melodic and upper tone pitch-matching were easier than lower tone pitch-matching. However, there were over two times as many errors when children were attempting to match the bottom note in a chord as compared to the top note. 52% of these errors involved the child singing a note higher than the prescribed pitch and 48% of the errors involved the child singing a note lower than the prescribed pitch. Therefore, there was no significant difference between the pitch direction the error was made. 72% of all errors made

involved the child singing the note in the chord, other than the one they were asked to reproduce. For this study, the researchers found that it became harder for children to pitch-match as intervals got smaller. It also seemed that the children improved with practice, as there were less errors made in the fourth middle tone pitch-matching test than in the first. Both the first and fourth middle tone pitch-matching tests were major triads and therefore should have been of similar difficulty level.

Based on these results, the researchers concluded that upper melody and harmony parts would be easier to learn than a lower harmony part for children between the ages of eight and eleven. It was also concluded that children of this age would have the most difficulty singing the middle harmony part in a song.

I feel that this study was carried out in such a way as to achieve quite accurate results. The test groups for age and sex were all equal, allowing for accurate comparisons to be made. I feel that the conclusions the researchers came to were logical based on the results of this study. As mentioned in the discussion section of this article, it would be interesting to use different test groups for this study. This test involved only children who were observed to be accurate singers by their teachers, and so it would be especially beneficial to repeat this test with less accurate singers to see if the results would still be the same. I think it would also be beneficial to repeat this test using a vocal model instead of a piano to give the stimulus for the children to respond to allow the children to sing more accurately, as was discussed earlier concerning vocal models. It would be interesting to see if this affected the results at all as this is a much debated topic of music educators.

Betty W. Atterbury and Lynn Silcox's (1993) research dealt with the effects of piano on singing accuracy. This test was done over the course of an entire school year and involved several research questions. The main purpose of the research was to determine whether there was a difference in the level of song acquisition between a group of kindergarten students taught to sing using harmonic piano accompaniment and a group that was taught without piano. Also asked, was whether there was a difference between the singing accuracy of students with high, moderate, and low music aptitude in each group over the course of a year, and whether there was a difference in music aptitude of students in each of the two groups.

The population used for this study was 205 kindergarten students from the Kindergarten Center in Graham, Maine in the 1990-1991 school year. For the purpose of this study, the kindergarten students were divided by class into a control group, which was taught to sing with harmonic piano accompaniment and an experimental group, which was taught to sing without the use of a piano. The control group was made up of eight kindergarten classes containing a total of 109 students, and the experimental group consisted of seven kindergarten classes, containing 96 students.

For the purpose of this study both groups were taught by the same teacher. The song *Pinto Pony* was used for this study. It is four phrases long, has an *aaba* form, and a range from D 1 to A1. All of the children were taught this song by rote within the first three weeks of school and then were asked to sing it individually into a tape recorder for the pre-test in the first week of October. Each child was then asked to sing this song again into the recorder in the first week of June. This repetition of the pre-test then acted as the post-test. Each child's performance was given a mark out of four for accuracy. The

Primary Measures Music Audiation test was given to each class in the last two weeks of May.

The results of this study revealed no significant difference between the accuracy of each group's ability to reproduce a whole song learned by rote. It was also discovered that the post-test scores for children with high musical aptitude were significantly higher than the scores of children with moderate or low musical aptitude. There was also no significant difference in the average Primary Measures Music Audiation scores between the experimental and control groups.

I found it very interesting that there were no significant difference between the control and experimental groups. I anticipated that the group trained without the use of a harmonic piano accompaniment would score higher on the post-test than the group trained with the use of piano accompaniment. I would tend to agree with the researchers in their hypothesis that part of the reason for the absence of significant difference between the singing accuracy of the two groups could be that there was not a very wide scale used when marking each singer. The children's performance was only marked out of four, not allowing for much differentiation between the abilities of some singers. Perhaps if there had been a wider scale, there would have been a larger difference between the singers in the two groups. One other possible difference is that the children were tested using a song they learned at the beginning of the year. It would be interesting to do another similar study using a song learned later in the year. This would perhaps give a more accurate picture of whether the children in the control group had become dependent on harmonic piano accompaniment in order to sing a song accurately, as they would have become accustomed to this throughout the year. Another interesting

repetition of this test would be to perform the test with several different grade levels. It would be interesting to note whether the use of harmonic accompaniment affected older children's singing accuracy differently than that of younger children or males differently than females. It would also be beneficial to do this test again with an equal number of kindergarten students in each class. This would allow the results to be more accurate as the control and experimental groups would be equal in size.

Catherine S. Jarjisian's (1983) research deals with the effect of different modes on song acquisition. Also examined in this study were the effects of socio-economic status and musical aptitude on these children's rote-singing abilities.

Jarjisian's study population was made up of grade one students from two different schools. One was a public school comprising primarily African American students, while the other was an Archdiocesan school consisting of primarily Caucasian students.

The students in the three grade one classes of each school were given the Primary Measure of Music Aptitude test prior to the commencement of the study. One of the school's music classes were taught by the students' regular teacher, while the other school's music classes were taught by the researcher. During their regular music classes, one of the classes at each school was trained for 10 to14 minutes using one of the following pitch-patterns: diatonic, pentatonic or a combination of both diatonic and pentatonic patterns. Each group was trained using pitch patterns and then was taught the same pentatonic and diatonic songs that would later be used in the study's test. Listening, movement and rhythmic activities were also incorporated into each group's classes. Following four weeks of instruction, each child individually sang two pentatonic and two diatonic songs.

The results of this study indicated that the groups of students taught using both pentatonic and diatonic pitch-pattern instruction scored much higher on the singing test than those students trained using just one of the modes of pitch patterns. There was also no significant difference found between the scores of those children trained using only pentatonic pitch patterns and those trained using only diatonic pitch patterns. As hypothesized by the researcher, children with higher scores on the music aptitude test scored slightly higher on the singing test than those with lower scores. The researcher also found no difference between the scores of the students from each school. Thus, the researcher concluded that the most effective method of instruction is a combination of pentatonic and diatonic pitch pattern instruction, regardless of the socio-economic status or the music aptitude of the students.

I was not surprised at the results of this study because a combination of pitch pattern instruction would allow for the child to be exposed to many different types of intervals. Then when presented with music of different modes, the child would have heard all of the intervals used before and would be better equipped to sing them. I do feel that the results of this test could have been more accurate if the number of students in each class at each school was equal. The study did not state how many children were in each class and therefore there is no way of knowing if the results were completely accurate. More accurate results could also have been reached regarding the socio-economic status question had the students all been taught by the same teacher. The teaching methods could have been different between the teachers for each school and therefore inaccurate conclusions could have been reached.

One factor affecting song acquisition that has been frequently researched is individual versus unison singing. The question asked in this research is whether it is easier for children to pitch-match and learn songs when singing alone or in a group. One such study was done by Mary Goetze (1985), which led to more research being done on this topic. Goetze's research compared singing accuracy when using text versus the syllable *loo*, as well as when singing individually versus in unison. For Goetze's study, a population of 165 students in kindergarten, grade one and three from three schools in Boulder and Denver, Colorado was used.

Goetze selected two songs of equal difficulty to be sung by the children. One song was to be performed individually, while the other was to be sung in unison. Both songs were sung to *loo* and with the text. The results of Goetze's study showed that the children consistently sang with more accuracy individually than in unison, whether singing with text or to *loo*. This was determined through the use of personal microphones recording each child's voice when singing in unison. It was also found that contour was less accurate when singing in unison, rather than individually. Results showed that kindergarteners sang more accurately in unison when singing to *loo* than when singing with text. Grade one students sang more accurately to *loo* individually, than in unison. With grade three subjects, the difference in accuracy between unison and solo singing was much less pronounced, regardless of whether singing to *loo* or with text; when singing individually, there was no difference between contour accuracy when singing to *loo* or with text. There was, however, a difference between kindergarten and grade three students in pitch and contour accuracy, as well as a significant difference between kindergarten and grade one in pitch accuracy. When singing in unison, there were

differences in contour accuracy between kindergarten and grade three, as well as between grades one and three. Overall, the best pitch-matching occurred when the subjects were singing individually with text. The largest difference in pitch accuracy between individual and unison singing was between kindergarten and grade one.

In regards to age, Goetze's results seem to show that pitch accuracy may develop at a younger age than accuracy in contour. As far as gender was concerned, boys sang less accurately than girls in all of the pitch test situations, with the difference increasing in unison singing. There was, however, no significant difference in contour accuracy between boys and girls in unison or individual singing.

From the results of this study, Goetze concluded that children may learn to sing more accurately if they are allowed to sing individually. Goetze suggested that if a child can sing accurately alone, he/she has learned pitch discrimination, how to self-monitor, how to remain attentive and how to properly use the vocal mechanism. She suggested that if a child sings inaccurately in unison, one of these necessities for accurate singing is not being achieved due to the presence of other voices singing simultaneously.

Nancy A. Cooper (1995) produced some interesting results, which differ from those obtained in Mary Goetze's study. The question being asked in Cooper's study was what effects gender, grade level, and individual versus unison singing have on vocal pitch accuracy in children grades one to five.

Cooper used 169 grades one to five children from a school in Akron, Ohio for her study. The children involved in this study usually had one 45-minute music class each week. For the purpose of this study, a four-beat melody with a range from C# 1to F# 1 was chosen. The children were asked to individually sing back the four-beat melody to

the syllable *loo*. The model used was a tape of a twelve-year old boy with an unchanged voice. In the test, the children were asked to echo back the melody by themselves and then also while the taped voice sang back the melody with them. This acted as the unison portion of the test. Half of the children did the individual test first, while the other half did the unison test first to ensure that the results were not biased based on which test was performed first.

The third grade was found to have the least accurate singing, while the fourth grade was found to be the most accurate. There was no significant difference found between individual and unison singing, although for the entire population, unison singing was slightly more accurate than individual singing. There was also no significant difference found between boys' and girls' accuracy in singing or between grades, except between grades three and four, where there was an improvement.

Cooper concluded that the reason for results that were contrary to prior research regarding individual versus unison singing could have been because of the melody's length. The melody used was only four beats long and therefore would have allowed for short-term memory advantages. Other suggested factors contributing to this difference were the use of a child model and the small range used. As was noted earlier in this essay, a child model has been suggested to be more accurate than some other models, such as a male changed voice and a female adult voice (Green 1990, 227). In addition to this, the range used for this test did not require a shift into the head voice and was also high enough not to tempt the children to switch out of their singing voice.

I would tend to agree with the Cooper's conclusion that the test results may not have been completely accurate because of the issues mentioned above. I would also agree

with Cooper's suggestion to repeat the test with one melody in the head register and one in the chest register and to use several melodic patterns rather than one. I felt that with the tests, short-term memory might have played a large role in the lack of difference between the individual and unison accuracy results. I would also suggest doing this test again using children from several different schools, rather than all from the same school. This would allow for more accurate results as the children would not all have the same music teacher, thus the familiarity with either individual or unison singing would not affect the results as much. More accurate conclusions would also result from using an equal number of males and females in each group, and an equal number of children in each class. Another factor affecting accuracy is that the unison singing portion of Cooper's test was performed using only one recorded voice that was singing with good pitch accuracy in addition to the child who was being tested. Discrepancies between Cooper and Goetze's results could have occurred due to this, as Goetze's test included several children at once in the unison singing portion.

Georgia A. Green's (1994) article also reveals different results than those found in the study by Goetze. The question being asked in Green's research was whether unison or individual singing has an effect on the vocal pitch accuracy of children in grades one, two, three, and five. Also studied was the difference in vocal pitch accuracy between grade levels and sexes.

Green's test population was 211 children from grades one, two, three, and five from the same inner-city elementary school. The song, *Bow Wow Wow* was taught to the children for use in this study. The range of this song is from D 1to the B 1. For the study, groups of eight children were brought into a separate room. Only four children were

individually recorded at a time, so four of the children were in the room simply to be present for the unison singing test. Each child was asked to sing the song individually in front of the group and then in unison with the other children in the group. Because of the 17 pitches in the piece, the researcher awarded a mark out of 33 in order to award a mark for each correct note sung and for each correct interval sung.

The results of this study indicated that the average mark for individual singing was 13.83, while the average mark for unison singing was 16.42. The average mark for female students was 16.89, while the average mark for males was 13.22. It was also discovered that there was improvement in both unison and individual singing as grade levels increased. The most noticeable increase in vocal pitch accuracy was in the unison singing marks between grades three and five.

I feel that Green's suggestion of trying this test again with more subjects present for the unison-singing test, was valid. With less children present for this study, it may have been easier for the children to hear themselves, thus allowing their unison singing marks to be higher than in the individual tests. I also agree that it would be beneficial to do the test again and have the children perform their individual singing tests in front of the researcher only, rather than in front of the other children. Especially in the higher grades, the students may be uncomfortable singing in front of their peers. This may have accounted for the small improvement in individual singing between grades three and five, and a significant improvement in the average unison singing marks for those same grades. I do not feel that I can comment on the accuracy of the conclusions reached regarding the grade level and gender comparisons of this study since it was not stated

whether the number of students per grade or the number of males versus females were equal.

III. PEDAGOGICAL IMPLICATIONS

This literature review has provided me with important information that I can use to develop better pedagogy for helping young children learn how to sing new song material. For example, it seems clear that the type and quality of vocal model can greatly affect the children's ability to match pitch. An instrument such as a child's unchanged voice or a female voice without vibrato is probably the best model to use. These models sound in the same octave as the children's voices and sound for a longer percentage of time on the desired pitch than a model with vibrato. Both of these factors seem to allow for greater ease and accuracy when children are attempting to match pitch.

The data in the study by Klinger, Campbell and Goolsby indicated that the approach to presenting a song to children needs to be determined based on the specific piece of music being taught. Although immersion is not always the most effective method of teaching a song, it does seem to be effective when a song is written in a question and answer form. The phrase-by-phrase method may be better suited to a larger piece of music that cannot be taught in its entirety at the same time or a piece of music not written in a question and answer form. This being said, I feel that a combination of the two approaches could be effective, with certain sections of a song being taught by immersion and others being taught using the phrase-by-phrase method.

Movement and visual aids appear to also be quite helpful when teaching song acquisition. These can allow the children to have a better understanding of the melody and its contour. Just as visual and kinesthetic aids can help children remember which

direction and how far apart the notes are, they can also be used to help build in practices such as articulation and dynamics. In addition to this, they can also distract children from possible tension-causing thoughts such as singing a high note. Research seems to be fairly inconclusive regarding the effects of gender on pitch-matching and song acquisition. Based on the various results from the studies above, further research needs to be done to discover if differences between boys and girls in pitch-matching and song acquisition are physiological or sociological.

When teaching a song in parts, it might be useful to consider which parts are easiest for children to hear and learn. The melody or top part is generally easier for children to reproduce accurately, with a solo melody being the easiest for young children to learn. In three-part songs, the bottom part is the next easiest for children to sing, with the middle part being the most difficult. This needs to be taken into consideration when deciding which part children are assigned to sing and also how much time is spent teaching each part. When choosing choral repertoire it also should be considered that smaller intervals are harder for children to reproduce accurately in parts. These chordal placement ideas may also influence repertoire choices with accompaniment. Whereas a song with an accompaniment that is higher than the vocal line or with chords that surround the melody may be more difficult for children to sing, a song with an accompaniment lower than the vocal line may be easier. The use of accompaniment should occur only after children are secure with the vocal parts that have been taught using an appropriate vocal model, such as a child’s voice or a female voice with little vibrato.

When teaching young children to sing, it seems to be most effective to use both pentatonic and diatonic modes, rather than focusing on just one. This makes sense as children will be better equipped to quickly learn a variety of songs because they will be comfortable singing both tones and semi-tones accurately. If taught using primarily one method, children will need to readjust quickly when presented with a song in another mode, allowing for inaccuracies to be present in their reproduction of the song.

Results seem to be contradictory when deciding whether to teach children to sing individually or in unison. For a choral conductor, unison singing is obviously used most frequently in a rehearsal setting. This alone is perhaps not the most effective approach to foster accurate singing. A solution to this would be to use unison singing in rehearsal, but to also allow time for children to sing individually or in smaller groups. This would allow for children to be able to hear themselves better and be able to self-correct, whereas in a choral setting this may not be done as easily. An approach to fitting individual singing into a choral rehearsal can be through singing games that alternate between individual and unison singing.

Children's song acquisition is a subject that needs to be considered by music educators and choral conductors alike. Much research has been done on this topic to attempt to arrive at the most effective methods for teaching songs to children. There are many factors that affect children's ability to learn a song, such as use of accompaniment, use of appropriate vocal models, teaching different modes, and individual versus unison singing. All of these factors must be considered when teaching song acquisition. As opinions are constantly changing on what is the most effective method of teaching song acquisition, more research could still be done. Areas of further research could include the

effects of gender stereotypes on song acquisition, the effects of immersion and phrase-by-phrase methods when teaching songs that are antecedent-consequent form, and whether pitch-matching is easier in a smaller rather than larger group.

IV. Reference List

Apfelstadt, Hilary. "Effects of Melodic Perception Instruction on Pitch Discrimination and Vocal Accuracy of Kindergarten Children." *Journal of Research in Music Education* 32 (1984): 15-24.

Atterbury, Betty W. and Lynn Silcox. "The Effect of Piano Accompaniment on Kindergartener's Developmental Singing Ability." *Journal of Research in Music Education* 41 (1993) 40-47.

Cooper, Nancy A. "Children's Singing Accuracy as a Function of Grade Level, Gender, and Individual versus Unison Singing." *Journal of Research in Music Education* 43 (1995): 222-231.

Green, Georgia A. "The Effect of Vocal Modeling on Pitch-Matching Accuracy of Elementary Schoolchildren." *Journal of Research in Music Education* 38 (1990): 225-231.

----------------------. "Unison versus Individual Singing and Elementary Student's Vocal Pitch Accuracy." *Journal of Research in Music Education* 42 (1994) 105-114.

Goetze, Mary. "Factors Affecting Accuracy in Children's Singing." PhD diss. University of Colorado at Boulder, 1985.

Gordon, Edwin E. "Developmental Music Aptitude as Measured by the Primary Measures of Music Audiation." *Psychology of Music* 7 (1979): 42-49.

Hornbach, Christina M. and Cynthia C. Taggart. "The Relationship Between Developmental Tonal Aptitude and Singing Achievement Among Kindergarten, First-, Second-, and Third-Grade Students." *Journal of Research in Music Education* 53 (2005): 322-331.

Jarjisian, Catherine S. "Pitch Pattern Accuracy and the Singing Achievement of Young Children." *Psychology of Music* 11 (1983) 19-25.

Klinger, Rita, Patricia Shehan Campbell, and Thomas Goolsby. "Approached to Children's Song Acquisition: Immersion and Phrase-by-Phrase." *Journal of Research in Music Education* 46 (1998): 24-34.

Liao, Mei-Ying. "The Effects of Gesture on Young Children's Pitch Accuracy for Singing Tonal Patterns." *International Journal of Music Education* 26 (2008): 197-209.

Moore, Randall S. "Effects of Age, Sex, and Melodic/Harmonic Patterns on Vocal Pitch-Matching Skills of Talented 8-11-Year-Olds." *Journal of Research in Music Education* 42 (1994) 5-13.

Yarbrough, Cornelia, Judy Bowers, and Wilma Benson. "The Effect of Vibrato on the Pitch-Matching Accuracy of Certain and Uncertain Singers." *Journal of Research in Music Education* 40 (1992): 30-38.